A Guide to Central Tendency

Choosing the Right Measure: Mean, Median, & Mode

What is Central Tendency?

Central tendency is a single value that attempts to describe a set of data by identifying the central position within that set.

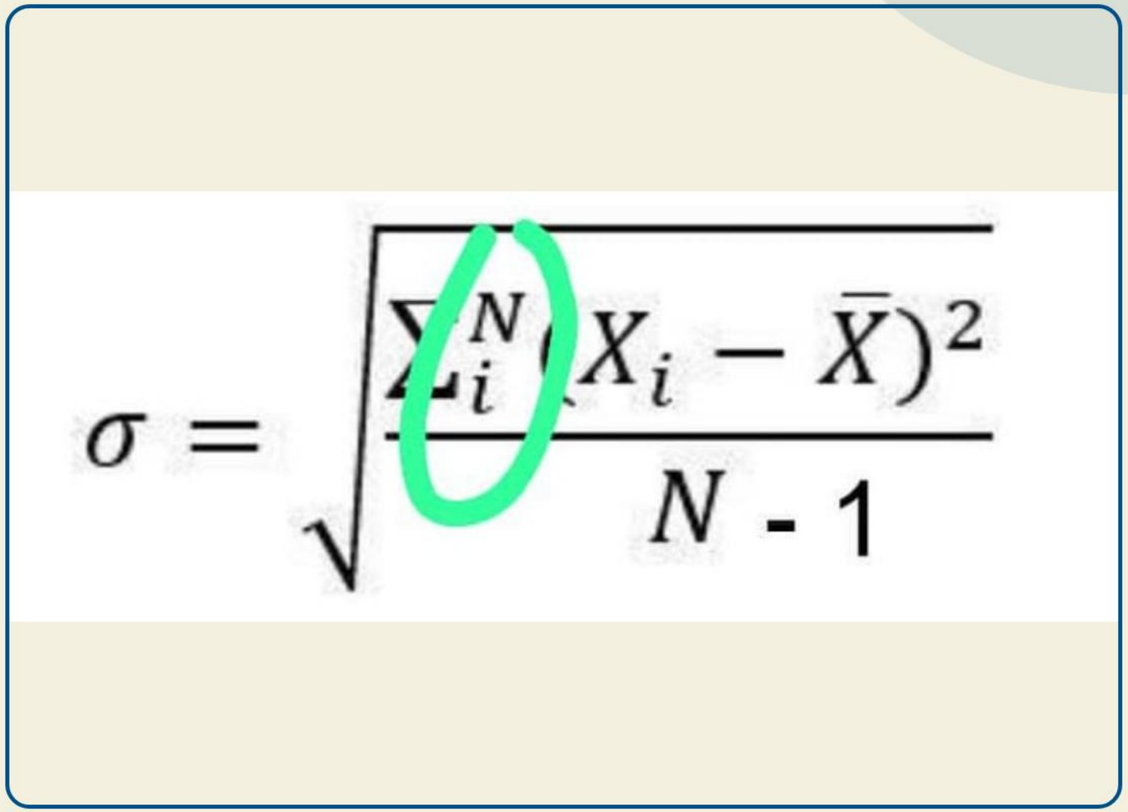
- It's a "summary statistic" that gives a quick overview of your data.
- It helps us find the "typical" or "most representative" value.
- The three most common measures are the Mean, Median, and Mode.

1. The Mean (Average)

What it is: The 'Average'

The mean is the most common measure. It is calculated by adding all the values and dividing by the number of values.

- **Pro:** It uses every single data point, making it a comprehensive measure.
- **Con:** It is highly sensitive to **outliers** (extremely high or low values), which can be misleading.

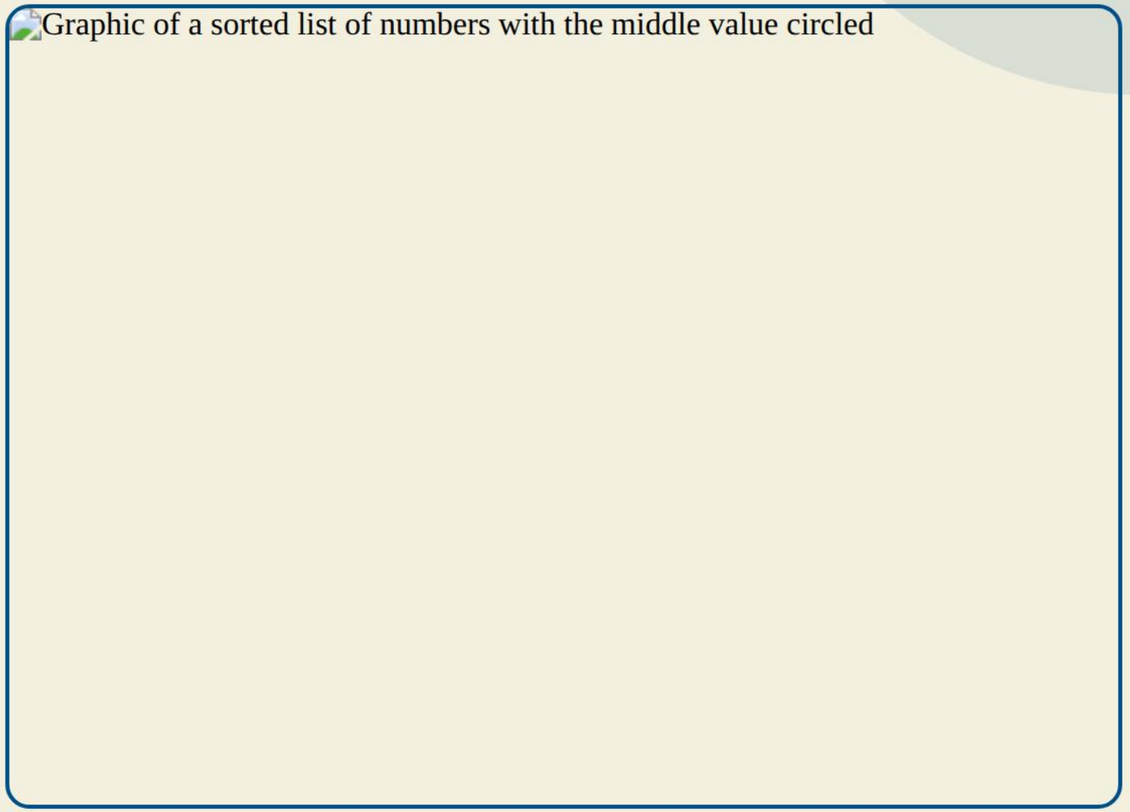


2. The Median (Middle)

What it is: The 'Middle' Value

The median is the value in the exact middle of a dataset that has been sorted from smallest to largest.

- **Pro:** It is **not affected by outliers**, making it a much better measure for skewed data (like income or house prices).
- **Con:** It ignores most of the data points, focusing only on the middle position.

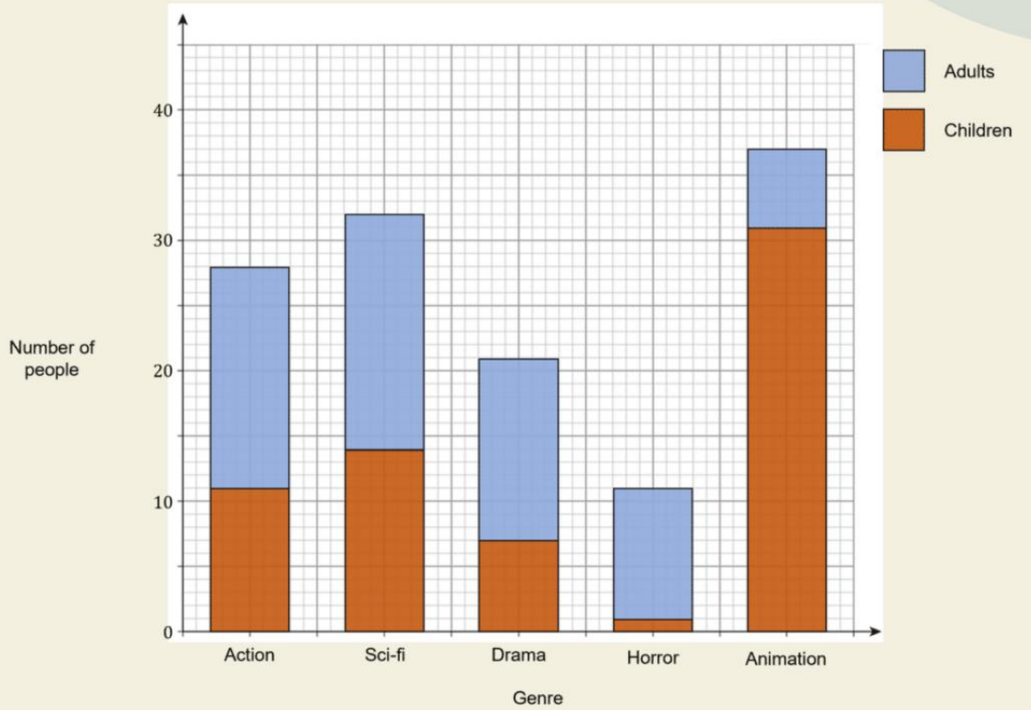


3. The Mode (Most Frequent)

What it is: The 'Most' Frequent

The mode is the value that appears most often in a dataset. A dataset can have one mode, multiple modes (bimodal), or no mode at all.

- **Pro:** It is the **only** measure that can be used for **categorical (nominal) data** (e.g., favorite color, t-shirt size).
- **Con:** It's not always unique and may not represent the center of the data well.



The Problem with Outliers

An Example: Cafe Salaries

Role	Annual Salary
Barista 1	$50,000
Barista 2	$52,000
Shift Lead	$55,000
Manager	$58,000
Owner's Nephew	**$200,000 (Outlier)**

How the Outlier Skews Data

The Mean

$83,000

This is misleadingly high. It doesn't represent a 'typical' salary at the cafe.

The Median

$55,000

This is the true 'middle' salary and is unaffected by the $200k outlier. It's a much better summary.

The Mode

None

Since no salary value repeats, there is no mode. This is also a valid (and useful) finding.

When to Use Which Measure?

Choosing the Right Tool for Your Data

Use the MEAN when...

Your data is symmetrical (a 'bell curve') and has no significant outliers.

Example: Test scores, daily temperatures, heights.

Use the MEDIAN when...

Your data is skewed or contains outliers.

Example: Income, wealth, housing prices.

Use the MODE when...

You have categorical (nominal) data, or you want to know the most popular choice.

Example: Favorite color, t-shirt size, survey answers.

Quick Comparison

Measure	What It Is	Best Data Type	Affected by Outliers?
Mean	The Average	Numeric (Interval/Ratio)	**Yes (Very)**
Median	The Middle Value	Numeric or Ordinal	**No**
Mode	The Most Frequent	Any (Best for Nominal)	**No**

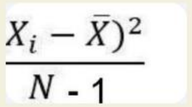
Questions?

Thank you for your attention.

Image Sources

https://static.vecteezy.com/system/resources/previews/024/790/205/non_2x/abstract-data-charts-statistic-graphs-finance-line-chart-and-marketing-histogram-graph-infographic-set-vector.jpg

Source: www.vecteezy.com

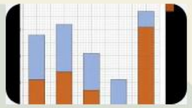


https://preview.redd.it/what-does-this-n-and-an-i-mean-in-this-formula-v0-rcm7uk0zi1de1.jpeg?width=640&crop=smart&auto=webp&s=6e7a0b503cf7d14004401ef628c8c667c611f221

Source: www.reddit.com

https://ecdn.teacherspayteachers.com/thumbitem/Circle-Number-Labels-1-30-Black-White-with-Neon-3289629-1657575150/original-3289629-1.jpg

Source: www.teacherspayteachers.com



https://mmerevise.co.uk/wp-content/uploads/2022/11/Q6-1024x660.png

Source: mmerevise.co.uk